Agrawal P

Data Engineer, 10+ years experience
Pune, India
https://augmntx.com/profile/data-engineer-agrawal-p/axf20126

+91 9820045154 | hr@augmntx.com

Data engineer with over 10 years of experince in Data migration, ETL pipelines and batch processing. Worked on designing data migration processes from RDMS system to Hadoop system and then on GCP cloud. Implemented python scripting for data cleansing and batch processing. Experience with Apach NiFi and other AWS services like AWS Glue, Lambda, S3.

INDUSTRIES: Healthcare, Telecommunication

TECHNICAL SKILLS

• SQL	10 years
• ETL	6 years
• Hadoop	5 years
• Big Data Technologies	5 years
• GCP	3 years
• Python	2 years
• Airflow	2 years
• AWS Glue	1 years
• AWS Lambda	1 years

SOFT SKILLS

Communication, Problem-solving, Teamwork

PROJECT DETAILS

• ETL Processing January-2020 - December-2021

Project Description: Implementation of ETL processes and Data Migration
Responsibilities: Providing solutions on customer datasets like Fraud Detection in Retail domain Improving data quality by performing DQ and DNTL checks Finding top 20 and bottom 20 search products by customers as per each state, department, and access_channel then generating report Python scripts for custom DAGs to pull data from S3 to GCS and load to Big Query Data cleaning using Python script Data migration from RDBMS to Hadoop and GCP Transforming existing SQL queries to Pig/Hive CI/CD pipelines using Jenkins
Technologies: Python, GCP, Hadoop
Industry: Healthcare

• Data Migration March-2014 - December-2019

Project Description: Design and implementation of data migration process from RDMS system to Hadoop data architecture
Responsibilities: Migrating 460 tables from oracle, MySQL and Netezza databases using Sqoop to Hadoop Developed utility to pull multiple tables from different environments by taking config parameters Creating layered architecture to bring data from incoming, staging, work and then to gold layer. Bring history then performing data cleansing activities like removing duplicates, removing extra characters, with strategy of loading full-load and incremental
Technologies: SQL, Hadoop, Scoop
Industry: Telecommunication

• ETL Pipeline and Data Migration

Project Description: Implementation of ETL pipeline and Data Migration from one system to another
Responsibilities: Leading a team of 5 developers for various projects Data migrations between different systems like Hadoop to GCP Big Query using DAGs File transfers using Apache Nifi Built a data processing pipeline on AWS cloud using Redshift, S3, Lambda, AWS Glue, Event rules Find the product search trends of the client and recommend them with specific offers Stakeholder management like Business Analysts and Product Leads for requirements and creating tech specs
Technologies: AWS, GCP, Nifi
Industry: Healthcare

EMPLOYMENT HISTORY

• Module Lead

Description: • Migrating 460 tables from oracle, MySQL and Netezza databases using Sqoop to Hadoop • Developed utility to pull multiple tables from different environments by taking config parameters • Creating layered architecture to bring data from incoming, staging, work and then to gold layer. • Bring history then performing data cleansing activities like removing duplicates, removing extra characters, with strategy of loading full-load and incremental

EDUCATION

• Bachelors degree in Computer Engineering from Amravati University 2008 - 2012

LANGUAGES

• English - advanced

END OF DOCUMENT
This PDF is automatically generated by AugmntX on 20 Nov 2026
Profile URL:https://augmntx.com/profile/data-engineer-agrawal-p/axf20126